



## King's Research Portal

DOI:

[10.1016/j.media.2014.05.010](https://doi.org/10.1016/j.media.2014.05.010)

*Document Version*

Early version, also known as pre-print

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Baumgartner, C. F., Kolbitsch, C., Balfour, D. R., Marsden, P. K., McClelland, J. R., Rueckert, D., & King, A. P. (2014). High-resolution dynamic MR imaging of the thorax for respiratory motion correction of PET using groupwise manifold alignment. *Medical Image Analysis*, 18(7), 939-952.  
<https://doi.org/10.1016/j.media.2014.05.010>

### Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### Take down policy

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# High-Resolution Dynamic MR Imaging of the Thorax for Respiratory Motion Correction of PET Using Groupwise Manifold Alignment

Christian F. Baumgartner<sup>a</sup>, Christoph Kolbitsch<sup>a</sup>, Daniel R. Balfour<sup>a</sup>, Paul K. Marsden<sup>a</sup>, Jamie R. McClelland<sup>b</sup>, Daniel Rueckert<sup>c</sup>, Andrew P. King<sup>a</sup>

<sup>a</sup>*Division of Imaging Sciences and Biomedical Engineering, King's College London,  
London, UK*

<sup>b</sup>*Centre for Medical Image Computing, University College London, London, UK*

<sup>c</sup>*Biomedical Image Analysis Group, Department of Computing, Imperial College London,  
London, UK*

---

## Abstract

Respiratory motion is a complicating factor in PET imaging as it leads to blurring of the reconstructed images which adversely affects disease diagnosis and staging. Existing motion correction techniques are often based on 1D navigators which cannot capture the inter- and intra-cycle variabilities that may occur in respiration. MR imaging is an attractive modality for estimating such motion more accurately, and the recent emergence of hybrid PET/MR systems allows the combination of the high molecular sensitivity of PET with the versatility of MR. However, current MR imaging techniques cannot achieve good image contrast inside the lungs in 3D. 2D slices, on the other hand, have excellent contrast properties inside the lungs due to the in-flow of previously unexcited blood, but lack the coverage of 3D volumes. In this work we propose an approach for the robust, navigator-less reconstruction of dynamic 3D volumes from 2D slice data. Our technique relies on the fact that data acquired at different slice positions have similar low-dimensional representations which can be extracted using manifold learning. By aligning these manifolds we are able to obtain accurate matchings of slices with regard to respiratory position. The approach naturally models all respiratory variabilities. We compare our method against two recently proposed MR slice stacking methods for the correction of PET data: a technique based on a 1D pencil beam navigator, and an image-based technique. On synthetic data with a known ground truth our proposed technique produces

significantly better reconstructions than all other examined techniques. On real data without a known ground truth the method gives the most plausible reconstructions and high consistency of reconstruction. Lastly, we demonstrate how our method can be applied for the respiratory motion correction of simulated PET/MR data.

*Keywords:* Respiratory motion correction, PET/MR motion correction, MR imaging of the thorax, Manifold learning, Manifold alignment

---

## 1. Introduction

Respiratory motion of the thorax is a complicating factor for many imaging techniques and image-based treatments. In radiation therapy it can cause healthy tissue to be irradiated (Seppenwoolde et al., 2002; von Siebenthal et al., 2007b), in image-guided interventions it causes misalignment between the static image-derived road map and the anatomy (King et al., 2009) and in positron emission tomography (PET) it leads to blurring of the scans, which in turn may lead to lower detectability of tumours (Zaidi and Del Guerra, 2011). In this paper we focus on the problem of PET imaging, specifically the use of magnetic resonance (MR) imaging to image the motion in a simultaneous PET/MR scanner.

### 1.1. The Problem of Motion in PET

One strategy for coping with respiratory motion in PET imaging is gating, a technique in which PET counts are only accepted when a respiration signal is within a small gating window. Respiration signals include external devices such as respiratory bellows (Klein et al., 1998) and camera-based systems (Nehmeh et al., 2002), or signals extracted from the PET data itself, for example by means of principal component analysis (PCA) (Thielemans et al., 2011). Gating methods reduce motion artefacts, but suffer from inefficient data usage, and the resulting images are of low signal-to-noise ratio (Würslin et al., 2013). More sophisticated techniques attempt to incorporate all counts into the reconstruction by using motion fields describing the respiratory motion to correct the PET data. The motion fields may either be used to retrospectively transform a number of reconstructed PET gates to a common reference in a process called reconstruct transform-average (RTA) (Dawood et al., 2006), or they may be incorporated directly into the PET reconstruction process for a motion corrected image-reconstruction (MCIR)

(Qiao et al., 2006). Motion fields can be extracted directly from the PET acquisitions (Bai and Brady, 2009; Dawood et al., 2006), however this approach may lack robustness for specific radiotracers with low background uptake (Würslin et al., 2013). Alternatively, motion fields may be derived from another modality such as computed tomography (CT). However, using CT comes at the cost of increased radiation exposure to the patient (Beyer et al., 2000). MR imaging is a promising alternative to CT because of its non-ionising nature, and its good soft tissue contrast. The recent emergence of hybrid PET/MR systems opens up the possibility of deriving the motion fields needed for PET motion correction from simultaneously acquired MR images (Tsoumpas et al., 2010; Würslin et al., 2013).

### *1.2. MR-Based Motion Correction*

Current MR imaging technology is limited by the lack of contrast in dynamic MR images of the lungs. The rapidly decaying MR signal from the lung parenchyma makes the lung appear dark in dynamic 3D images (Robson et al., 2003), prohibiting robust motion field estimation. Since lung motion is not homogeneous, accurate motion fields cannot be simply interpolated between the lung borders (Ding et al., 2009). Furthermore, dynamic 3D MR images suffer from low image resolution and relatively long acquisition times, which can lead to motion blurring, further limiting the accuracy of motion estimation. Dynamic 2D scans, on the other hand, can be acquired in a shorter time frame, have excellent in-plane resolution and, when acquired at systole, have high contrast in the lungs due to the in-flow of previously unexcited blood into the slice. However, they lack the coverage of 3D scans. Examples of dynamic 2D and 3D acquisitions are shown in Figure 1.

It has been recognised that combining the favourable contrast and resolution properties of 2D slices with the coverage of 3D slices is possible by acquiring 2D slices multiple times in a slice-by-slice fashion and then retrospectively reconstructing dynamic 3D volumes at each time point using a slice stacking technique. Dikaio et al. (2012) compared 2D slice-by-slice and 3D MR acquisition protocols for motion correcting PET/MR data on synthetic data generated from volunteer MR scans, concluding that motion estimation from dynamic 3D volumes lacks accuracy. In order to reconstruct 3D volumes from the 2D data the authors used a simple slice stacking technique based on image similarities of neighbouring slice positions. Würslin et al. (2013) used a similar slice-by-slice acquisition protocol, but in addition to the slice data, the authors acquired a 1D navigator echo, also known as a

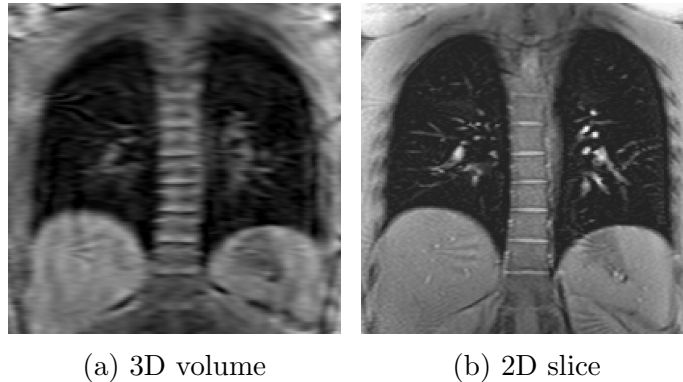


Figure 1: Comparison of (a) a coronal slice through a dynamic 3D MR volume; (b) a coronal dynamic 2D MR slice acquired from a similar anatomical position. The 2D slice has much improved contrast inside the lungs due to the in-flow of unexcited blood.

pencil beam navigator (Ehman and Felmlee, 1989), which they used to gate the 2D slices to reconstruct dynamic 3D volumes. The resulting volumes were used to correct clinical PET scans for respiratory motion.

It is well known that respiratory motion exhibits non-negligible variations between and within breathing cycles such as amplitude variations, baseline shifts and hysteresis (Blackall et al., 2006; von Siebenthal et al., 2007b; McClelland et al., 2011). McClelland et al. (2013) classified these variations into two categories: inter-cycle variation, i.e. the motion path during one breathing cycle is different from the motion path of a different breathing cycle, and intra-cycle variation, i.e. the motion path during inspiration is different from the one during expiration. A simple 1D navigator, as was used in Würslin et al. (2013), cannot accurately capture these variations. This effect is aggravated by the fact that PET typically involves acquisitions over tens of minutes, a time-frame in which many patients may change their respiration pattern, e.g. because of varying degrees of relaxation whilst in the scanner.

An alternative approach, that has the potential to cope with such variations, was proposed by King et al. (2012). In this work, a motion model based on low-resolution 3D MR dynamics was built before the PET acquisition, and then applied using fast 2D MR slices during the acquisition. Using this method motion fields were estimated for any time-point of the PET acquisition, and used to correct the PET data. However, the authors pointed out that motion fields derived from low-resolution 3D volumes restrict the

accuracy of such a technique inside in the lungs.

A slice stacking technique that could capture breathing variation was proposed by von Siebenthal et al. (2007a). The technique was based on an interleaved slice acquisition protocol, where a 2D navigator slice at a fixed slice position was acquired before and after data slices at varying slice positions. This approach was hampered by a low scan efficiency (only half of the data acquired was actually used for reconstruction). Furthermore, it required fast non-cardiac gated acquisitions over tens of minutes, which makes the approach infeasible in a PET/MR scenario where it is desirable to simultaneously perform other scans of clinical relevance.

In this paper we present a novel navigator-less slice stacking method for retrospectively reconstructing 3D dynamic MR volumes from 2D slices, which is based on the simultaneous embedding and groupwise alignment of underlying low-dimensional manifold embeddings of data acquired at different slice positions. The method consists of three main steps: We first acquire data from different slice positions in a slice-by-slice protocol (similar to that used by Dikaio et al. (2012) and Würslin et al. (2013)). We then find respiratory correspondences between data from different slice positions using manifold alignment. Lastly, we reconstruct dynamic 3D+t volumes from the 2D data based on the previously found respiratory correspondences. An overview of the method is shown in Figure 2. The proposed technique is entirely image based (i.e. requires no additional navigator), has 100% data efficiency and has the potential to inherently model all intra-cycle and inter-cycle variabilities. From the reconstructed volumes robust motion fields in the thorax (including the lungs) can be obtained, which can be used to motion correct simultaneously acquired PET images.

### *1.3. Manifold Alignment*

Manifold learning is a powerful tool for non-linear dimensionality reduction of complex high-dimensional data and various manifold learning algorithms have been proposed. In recent years manifold learning was shown to be useful in the analysis of motion in medical images, making use of the fact that similar points in the low-dimensional space correspond to similar motion states. Applications include the region-wise separation of cardiac and respiratory motion (Bhatia et al., 2012), retrospective reconstruction of respiratory-gated lung CT volumes (Georg et al., 2008) and extraction of respiratory gating navigators from MR and ultrasound images (Wachinger et al., 2011).

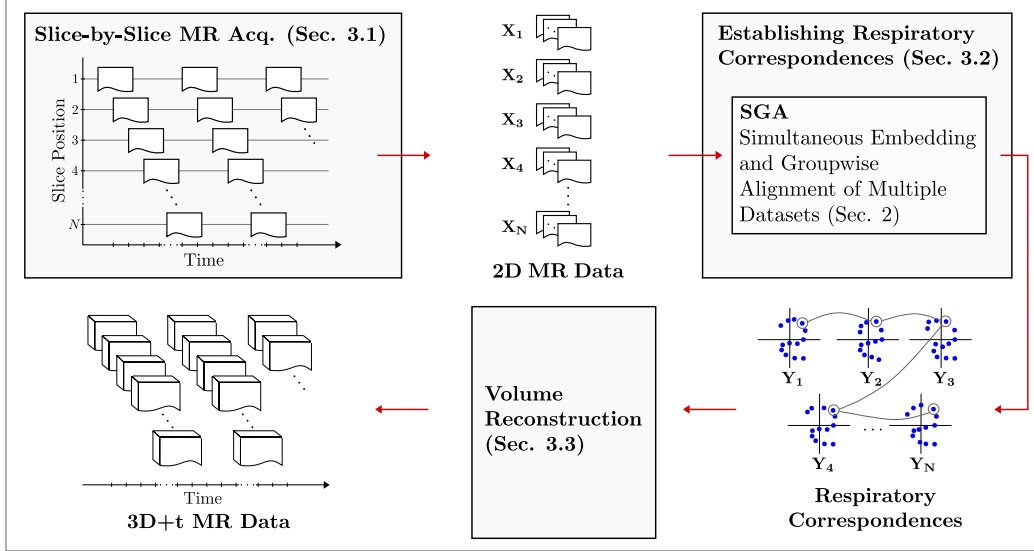


Figure 2: Overview of the proposed method. First, 2D MR data is acquired in a slice-by-slice fashion. Respiratory correspondences between different anatomical positions are established by means of simultaneous groupwise manifold alignment (SGA) of data from different slice positions. The  $\mathbf{X}_p$  denote data acquired at slice position  $p$ , the  $\mathbf{Y}_p$  denote low-dimensional manifold embeddings thereof. Lastly, for each 2D input slice, a 3D volume is reconstructed based on the established correspondences.

Manifold alignment can be used to establish correspondences between multiple related datasets, which are not directly comparable in high-dimensional space, but have a similar low-dimensional manifold structure. There are two general approaches to manifold alignment: 1) The datasets are embedded in a common low-dimensional space in a single *simultaneous* embedding, either with prior knowledge of corresponding points (Ham et al., 2005; Zhai et al., 2010), or without such knowledge (Torki et al., 2010; Baumgartner et al., 2013); 2) The datasets are embedded *separately*, and are then transformed to the same coordinate system in a subsequent alignment step either with known correspondences using a shape matching technique like Procrustes analysis (Wang and Mahadevan, 2008), or without known correspondences using normalisation (Georg et al., 2008).

In this paper we propose a novel scheme for manifold alignment through simultaneous embedding and apply it to the problem of retrospective recon-

struction of dynamic 3D MR volumes from 2D slices. In particular, we take advantage of the fact that data acquired at neighbouring slice positions lie on manifolds with similar structure to establish correspondences in respiratory state between those slice positions. Our novel manifold alignment scheme is an extension of our previous work (Baumgartner et al., 2013), extended to use an improved inter-dataset kernel, which accounts for anatomical differences between slice positions. We also include a more thorough validation of our technique, comparing it to the methods proposed by Dikaïos et al. (2012) and Würslin et al. (2013), and validate it on synthetic PET data derived from real 4D CT scans.

The paper is structured as follows: In Section 2 we give a brief review of the theory of manifold learning and simultaneous groupwise manifold alignment (SGA). In Section 3 we show how SGA can be used to perform reconstruction by slice stacking of 2D slice-by-slice acquisitions. In particular, in Section 3.1 we explain the slice-by-slice acquisition in detail, in Section 3.2 we explain how the SGA theory can be applied to establish respiratory position correspondences between different slice positions and in Section 3.3 we show how volumes can be stacked based on those correspondences. Section 4 describes the experiments performed, and the results are presented in Section 5. Lastly, Sections 6 and 7 contain the discussion and conclusion.

## 2. Theory

Manifold alignment by means of simultaneous embedding of multiple datasets is the core element of the proposed slice stacking technique. In this section, we first provide some background on manifold learning and alignment, and then outline the theory behind our simultaneous groupwise alignment scheme.

### 2.1. Manifold Learning

In natural datasets the dimensionality is often artificially high. Consider, for example, a set of coronal slices such as the one shown in Figure 1b, which differ from one another by a deformation of tissue due to respiratory motion. The images lie in the very high-dimensional image space  $\mathbb{R}^D$ , where  $D$  is the number of pixels per image. However, the images vary due to a much smaller number of degrees of freedom and therefore can be viewed as a set of points



on a manifold of many fewer dimensions  $\mathbb{R}^d : d \ll D$ , embedded in high-dimensional space. This embedding can be uncovered using dimensionality reduction techniques.

In addition to classical linear dimensionality reduction techniques such as PCA, a large number of non-linear dimensionality reduction techniques such as locally linear embeddings (LLE) (Roweis and Saul, 2000), Laplacian Eigenmaps (LEM) (Belkin and Niyogi, 2003) and Isomap (Tenenbaum et al., 2000) have been proposed in recent years. These techniques are collectively known as manifold learning.

## 2.2. Simultaneous Embedding of Two Datasets

In some applications multiple datasets may not be directly comparable in high-dimensional space, but may nevertheless lie on similar manifolds. For example, this is the case when the datasets are slices acquired at two different slice positions, which are deformed by the same respiratory mechanics. Points close to each other on one embedded manifold correspond to similar respiratory positions. However, this is not true for points from embeddings obtained from two different datasets. In this case correspondences between datasets can be established in the low-dimensional space by means of manifold alignment.

Given two such high-dimensional datasets  $\mathbf{X}_1, \mathbf{X}_2$  (e.g. data from two slice positions), for each element  $\vec{X}_i^{(1)}, \vec{X}_j^{(2)} \in \mathbb{R}^D$  (e.g. single slices from these two slice positions), the aligned embeddings  $\mathbf{Y}_1, \mathbf{Y}_2$  with elements  $\vec{Y}_i^{(1)}, \vec{Y}_j^{(2)} \in \mathbb{R}^d : d \ll D$ , can be found by minimising the total embedding error

$$\phi_{tot}(\mathbf{Y}_1, \mathbf{Y}_2) = \phi_1(\mathbf{Y}_1) + \phi_2(\mathbf{Y}_2) + \mu \cdot \phi_{12}(\mathbf{Y}_1, \mathbf{Y}_2), \quad (1)$$

where  $\phi_1, \phi_2$ , are the *intra-dataset* embedding errors and  $\phi_{12}$  is the *inter-dataset* embedding error. The weighting parameter  $\mu$  regulates the influence of the inter-dataset error term. Increasing  $\mu$  forces the embeddings to be closer together, but setting it too high may distort the natural manifold structure of the data.

In this paper we derive the intra-dataset error terms from LLE. LLE tries to preserve locally linear relations of the high-dimensional data in the low-dimensional embeddings. It is assumed that each high-dimensional point can be reasonably well reconstructed using a linear combination of its  $k$  nearest neighbours. The optimal reconstruction weights  $W_{ij}$  for each point  $i$  from its respective neighbours  $j$  can be calculated in closed form as described in

Roweis and Saul (2000). The intra-dataset embedding errors,  $\phi_1$ , and  $\phi_2$ , which preserve the local relations of the high-dimensional data can then be expressed as

$$\phi_m(\mathbf{Y}_m) = \sum_i \left| \vec{Y}_i^{(m)} - \sum_{j \in \eta(i)} W_{ij}^{(m)} \vec{Y}_j^{(m)} \right|^2, \quad m \in \{1, 2\}, \quad (2)$$

where for each dataset  $\eta(i)$  is the respective neighbourhood of data point  $i$ .

For the inter-dataset error a different cost function is used. The embedding error of  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$  is defined as

$$\phi_{12}(\mathbf{Y}_1, \mathbf{Y}_2) = \sum_{i,j} \left( \vec{Y}_i^{(1)} - \vec{Y}_j^{(2)} \right)^2 U_{ij}, \quad (3)$$

where  $U_{ij} = K(\vec{X}_i^{(1)}, \vec{X}_j^{(2)})$  is a (non-symmetric) similarity kernel. For high similarity values  $U_{ij}$  the error can only be minimised if  $\vec{Y}_i^{(1)}$  and  $\vec{Y}_j^{(2)}$  are close in the simultaneous embedding. The choice of this kernel is non-trivial and is a key element of our proposed method. In the case of a set of *a priori* known correspondences it may consist of fixed connections (Ham et al., 2005; Zhai et al., 2010). In the case of no known correspondences it may be based on similarities in high-dimensional space (Torki et al., 2010). Our choice of the kernel will be discussed in Section 3.2.

Independent of the kernel choice the minimisation of the whole cost function  $\phi_{tot}$  can be rewritten in a single matrix expression,

$$\operatorname{argmin}_{\mathbf{Y}_1, \mathbf{Y}_2} Tr \left( \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix}^T \begin{bmatrix} \mathbf{M}_1 + \mu \mathbf{D}_1 & -\mu \mathbf{U} \\ -\mu \mathbf{U}^T & \mathbf{M}_2 + \mu \mathbf{D}_2 \end{bmatrix} \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \right), \quad (4)$$

where  $\mathbf{D}_1$  and  $\mathbf{D}_2$  are the row and column sums of  $\mathbf{U}$  as diagonal matrices, i.e.  $D_{ii}^{(1)} = \sum_j U_{ij}$  and  $D_{jj}^{(2)} = \sum_i U_{ij}$ ,  $\mathbf{M}_1$  and  $\mathbf{M}_2$ , are the recentred reconstruction weight matrices  $\mathbf{M}_m = (\mathbf{I} - \mathbf{W}_m)^T (\mathbf{I} - \mathbf{W}_m)$ , and  $Tr(\cdot)$  is the trace operator. This problem now has the same form as the standard LLE embedding (Roweis and Saul, 2000), that is,

$$\operatorname{argmin}_{\mathbf{V}} Tr(\mathbf{V}^T \mathbf{L} \mathbf{V}), \quad (5)$$

where  $\mathbf{L}$  is the augmented matrix from Eq. (4) and  $\mathbf{V}$  are the augmented embeddings. Under the constraint that  $\mathbf{V}^T \mathbf{V} = \mathbf{I}$ , the simultaneous aligned embedding is given by the second smallest to the  $(d+1)$ -th smallest eigenvectors of  $\mathbf{L}$ .

### 2.3. Simultaneous Groupwise Embedding of Multiple Datasets

We now describe our simultaneous groupwise manifold alignment scheme in the context of this theory. This technique was first presented with preliminary evaluation in Baumgartner et al. (2013).

Eq. (4) can be easily extended to three or more manifolds by further augmenting  $\mathbf{V}$  and  $\mathbf{L}$ , as was described e.g. in Torki et al. (2010). However, in Baumgartner et al. (2013) we showed that this approach is not optimal, for two main reasons: 1) The problem becomes increasingly unstable when increasing the number of datasets  $\mathbf{X}_{\mathbf{p}}$  in the cost function. That is to say, the embeddings tend to collapse onto a single point or line, or not to be aligned at all, depending on the choice of the parameter  $\mu$  in Eq. (1). For around 30 datasets, which is the number of datasets (i.e. slice positions) we are embedding in this work, it is not possible to select a suitable  $\mu$ ; 2) The manifold structure may not be similar enough across all datasets to justify a simultaneous embedding of all of them.

Instead we propose to embed the datasets simultaneously in overlapping groups of two, producing a much more stable problem. For  $N$  high-dimensional inputs  $\mathbf{X}_1, \dots, \mathbf{X}_N$ , the datasets are embedded in  $N - 1$  groups  $\mathcal{G}^{(1)}, \dots, \mathcal{G}^{(N-1)}$ . In Section 3.2 we will define  $\mathbf{X}_1$  as all slices acquired at slice position 1,  $\mathbf{X}_2$  as all slices acquired at slice position 2 and so on. But the theory is generally applicable to other types of data as well. Each group  $\mathcal{G}^{(p)}$  contains the simultaneous embeddings of two datasets  $\mathbf{X}_{\mathbf{p}}$ , and  $\mathbf{X}_{\mathbf{p}+1}$ , i.e.  $\mathcal{G}_1^{(p)} = \mathbf{Y}_{\mathbf{p}}$ , and  $\mathcal{G}_2^{(p)} = \mathbf{Y}_{\mathbf{p}+1}$ . The groups are overlapping in the sense that  $\mathcal{G}_2^{(p)} = \mathcal{G}_1^{(p+1)}$ . This means that each dataset  $\mathbf{X}_{\mathbf{p}}$  (except  $\mathbf{X}_1$  and  $\mathbf{X}_N$ ) is embedded in two distinct groups  $\mathcal{G}^{(p)}$  and  $\mathcal{G}^{(p-1)}$  or, in other words, both  $\mathcal{G}^{(p)}$  and  $\mathcal{G}^{(p+1)}$  contain an embedding of the dataset  $\mathbf{X}_{\mathbf{p}}$ .  $\mathbf{X}_1$  and  $\mathbf{X}_N$  are embedded in only one group because there is no more data before or after them, respectively. Figure 3 shows an artificial example of a groupwise embedding in  $d = 2$  dimensions and the relations between the groups.

By embedding the datasets in overlapping groups in this manner correspondences between the different datasets can be found by going from group to group. The two members of each group are aligned due to the simultaneous embedding, and the connections to the next group are deterministically known through the group overlap, i.e. because of the dataset that the groups share. For example, consider an arbitrary point  $\vec{Y}_i^{(p)}$  (labelled with a square in Figure 3), which is embedded in the first member manifold of group  $\mathcal{G}^{(p)}$ , i.e. in  $\mathcal{G}_1^{(p)} = \mathbf{Y}_{\mathbf{p}}$ . Since within the group the manifolds are aligned, the

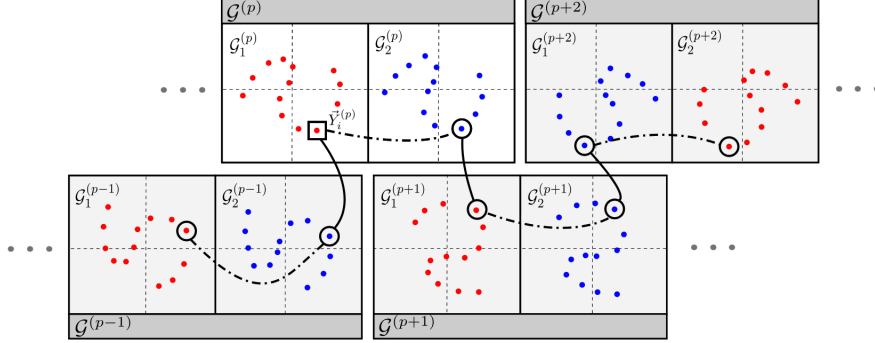


Figure 3: Schematic illustration of groupwise manifold alignment. The curved lines illustrate the manifold connections through the group overlap (solid), or through aligned embedding (dotted).

closest neighbour on  $\mathbf{Y}_{\mathbf{p}+1} = \mathcal{G}_2^{(p)}$  can be found directly (see dotted lines in Figure 3).  $\vec{Y}_i^{(p)}$  and its corresponding point on  $\mathbf{Y}_{\mathbf{p}+1}$  can be looked up directly in the neighbouring groups  $\mathcal{G}^{(p+1)}$ , and  $\mathcal{G}^{(p-1)}$ , since  $\mathcal{G}_2^{(p-1)} = \mathcal{G}_1^{(p)}$  and  $\mathcal{G}_1^{(p+1)} = \mathcal{G}_2^{(p)}$  (see solid lines in Figure 3). In this manner the low-dimensional points closest to  $\vec{Y}_i^{(p)}$  on all other manifolds embeddings can be found iteratively.

Note that by choosing the closest neighbour within the groups (dotted lines) and transporting that between groups (solid lines) instead of the actual point coordinates, we incur small errors in the process of iteratively looking up correspondences to a low-dimensional point. In order to propagate the actual point coordinates we investigated interpolating the location of the point in the new group by looking at the  $k$  nearest neighbours on the current manifold embedding and finding an average of those points in the manifold embedding of the new group. However, in our application of volume reconstructions, this only slightly influenced the results and therefore, for simplicity, we chose to transport just the closest neighbour as described above.

### 3. Materials and Methods

We now describe our specific implementation of the theory outlined in the last section.

The proposed technique can be divided into three steps: Slice-by-slice

data acquisition (Section 3.1), establishing correspondences in respiratory position between slices acquired at different slice positions based on simultaneous groupwise embedding of multiple datasets (Section 3.2), and dynamic 3D reconstruction by means of slice stacking (Section 3.3).

### *3.1. Slice-by-Slice Acquisition*

We imaged the volume of interest by sequentially acquiring coronal slices at shifting slice positions covering the whole region of interest, as illustrated in Figure 2. In order to sufficiently sample all respiratory states each slice position was acquired 50 times. The acquisition time for each slice was 160 ms. To maximise vessel contrast, and to minimise cardiac motion, only one slice was acquired per heartbeat at systole resulting in a temporal resolution of approximately one slice per second. Note that in a simultaneous PET/MR scenario the rest of the cardiac cycle could be used for other clinical scanning.

The acquisitions were carried out on a Phillips Achieva 3T MR scanner using a T1-weighted gradient echo sequence with an acquired in-plane image resolution of  $1.4 \times 1.4 \text{ mm}^2$  and a slice thickness of 8 mm. To cover a region of interest in the thorax including most of the lungs in anterior-posterior (A-P) direction typically 29-34 slice positions were needed, which resulted in a total acquisition time of 24-28 minutes. In order to decrease the effective slice thickness, and to help to establish respiratory correspondences between neighbouring slice positions later on, the slice positions were defined in an overlapping fashion with a slice overlap of 4 mm. This resulted in an effective slice thickness of 4 mm in a reconstructed volume. To avoid polarisation artefacts from previous slices, an acquisition scheme was used in which the slice position number increased with a step size equal to the rounded square root of the total number of slice positions.

### *3.2. Establishing Correspondences of Respiratory Positions*

In order to reconstruct volumes from the 2D slices, correspondences in respiratory position across the slice positions must be established. We accomplished this using the simultaneous groupwise manifold alignment (SGA) technique introduced in Section 2.3.

Slices acquired at each of the slice positions were defined as the high-dimensional input datasets  $\mathbf{X}_{\mathbf{p}}$ , where  $p$  is the slice position. Since all slice positions undergo approximately the same transformations due to respiratory motion they lie on similar low-dimensional manifolds  $\mathbf{Y}_{\mathbf{p}}$  and points that are close together on this manifold indicate similar respiratory positions.

The alignment of these embeddings allowed the necessary correspondences in respiratory position to be established.

One key design choice is the choice of the inter-dataset kernel  $K(\vec{X}_i^{(p)}, \vec{X}_j^{(q)})$  from Equation (3) for the slice positions  $p$  and  $q$ . To facilitate this choice the groups were naturally chosen such that neighbouring slice positions form groups, though other methods of forming groups are also conceivable and will be investigated in future work. Some anatomical similarities can be expected from neighbouring slices on which a similarity measure can be based. The fact that neighbouring slices are in fact overlapping (as described in Section 3.1), and hence share some anatomical features, further facilitates this task. Two kernel choices were investigated, as discussed below.

### 3.2.1. Image Similarity-based Inter-dataset Kernel

In Baumgartner et al. (2013) we proposed an inter-dataset kernel based directly on the image similarity of neighbouring slices. The kernel  $\mathbf{U}^{(\mathbf{p}, \mathbf{q})}$  relating the slices from slice position  $p$  to the slices from slice position  $q$  was defined as

$$U_{ij}^{(p,q)} = K(\vec{X}_i^{(p)}, \vec{X}_j^{(q)}) = \exp\left(-\frac{\tilde{\mathcal{L}}_2(\vec{X}_i^{(p)}, \vec{X}_j^{(q)})^2}{2\sigma^2}\right), \quad (6)$$

where  $\tilde{\mathcal{L}}_2$  denotes the normalised  $\mathcal{L}_2$ -distance, which is a scaled version of the  $\mathcal{L}_2$ -distance such that the maximum  $\tilde{\mathcal{L}}_2$  value between slices acquired at two neighbouring slice positions is equal to 1. The parameter  $\sigma$  governs the kernel shape. The normalisation step allows the value of  $\sigma$  to be subject independent. Normalised cross correlation and mutual information were also investigated as possible distance measures, but in our experiments  $\tilde{\mathcal{L}}_2$  performed the best. For the remainder of this paper we will refer to SGA using this similarity-based kernel as SGA.SIM.

### 3.2.2. Registration-based Inter-dataset Kernel

Even though the overlapping slices used in this study cover some common anatomy and consequently look similar, the small changes in slice position between two neighbouring slices still cause non-negligible differences in the images. In particular, a shift in slice position might cause a similar apparent deformation to a change in respiratory position (e.g. a shift in the diaphragm position and/or a change in lung area). Hence, basing the inter-dataset

kernel directly on the image similarities may adversely affect the resulting reconstructions.

To address this problem, in this work we propose a novel registration-based inter-dataset kernel, which incorporates knowledge of the approximate relations between adjacent slice positions. Those relations were obtained using registration of exhale slices, which were subsequently transported to different respiratory states using transformations obtained from a second set of registrations. In this manner approximations of the relations between neighbouring slice positions at any respiratory position were found, which were then incorporated in the similarity kernel in order to improve the similarity measure between slice positions. This process is explained in more detail below. The relations between the slices and transformations that we will refer to in the following are shown in Figure 4.

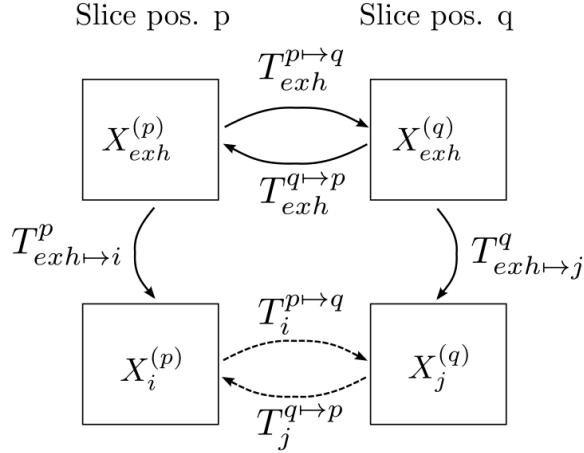


Figure 4: Relations between the slices at the neighbouring slice positions  $p$  and  $q$  and the different transformations between them. The transformations  $T_{exh}^{p \mapsto q}$ , and  $T_{exh}^{q \mapsto p}$  between two exhale slices  $X_{exh}^p$ , and  $X_{exh}^q$  were transported to different respiratory positions using the transformations  $T_{exh \mapsto i}^p$ , and  $T_{exh \mapsto j}^q$ , to arrive at approximate transformations  $T_i^{p \mapsto q}$ , and  $T_j^{q \mapsto p}$  between two slices  $X_i^p$ , and  $X_j^q$  at arbitrary time points  $i$ , and  $j$ .

In a first step an exhale volume consisting of exhale slices  $X_{exh}^p$  was reconstructed from the slice-by-slice data. Because the exhale state is very reproducible (Blackall et al., 2006) this can be accomplished by simply stacking the slices with the smallest lung area, which can be found reliably by calculating the mean intensity of each slice position; the slices with the highest mean

intensity corresponded to the smallest lung area. Note that no additional imaging was necessary for this step.

The relations between neighbouring exhale slices at slice positions  $p$  and  $q$  are captured by the transformation  $T_{exh}^{p \rightarrow q}$  which was obtained by registering  $X_{exh}^p$  to its neighbour  $X_{exh}^q$  using B-spline registration (Rueckert et al., 1999) with a control point spacing of 2.8 mm, a smoothness penalty term of  $\lambda = 0.01$  and sum of squared differences (SSD) as similarity measure. A relatively small control point spacing was chosen because changes from slice position to slice position typically do not involve large displacements. The corresponding motion fields  $D_{exh}^{p \rightarrow q}$  can be obtained by subtracting the grid locations from the transformed grid points, i.e.

$$D_{exh}^{p \rightarrow q}(x', y') = T_{exh}^{p \rightarrow q}(x', y') - (x', y'), \quad (7)$$

where  $(x', y')$  are the grid points of the transformation.

In order to make this mapping across slice positions available at all respiratory positions, the transformations  $T_{exh}^{p \rightarrow q}$  were transported to the different respiratory states. This can be viewed as transporting the motion fields from the exhale coordinate system to the coordinate systems of different respiratory states which are non-rigidly deformed with respect to the exhale coordinate system. We adapted the approach described by Rao et al. (2002) who used this concept to transport motion fields from one patient coordinate system to another. To obtain mappings to different respiratory states, we performed additional registrations within the slice positions to different time points. That is, we computed the transformations  $T_{exh \rightarrow i}^{(p)}$  by registering  $X_{exh}^p$  to  $X_i^p$ . For this step, B-spline registration with a larger control point spacing of 14 mm was used, because the deformations due to respiratory motion are larger than the deformations from slice position to slice position.

To obtain  $T_i^{p \rightarrow q}$  we transported the motion fields  $D_{exh}^{p \rightarrow q}$  from the coordinate system  $(x', y')$  to the coordinate system of respiratory state  $i$ , which we denote by  $(x, y)$ . If the motion vector at positional coordinates  $(x'_0, y'_0)$  in the coordinate system of  $exh$  is equal to  $d_{exh}^{p \rightarrow q}$ , the transported vector at the location  $(x_0, y_0)$  and respiratory position  $i$  is given by  $d_i^{p \rightarrow q}$ , where

$$(x_0, y_0) = T_{exh \rightarrow i}^{(p)}(x'_0, y'_0) \quad \text{and} \quad d_i^{p \rightarrow q} = J^{-1} \cdot d_{exh}^{p \rightarrow q}. \quad (8)$$

Here,  $J$  is the Jacobian matrix of  $(T_{exh \rightarrow i}^{(p)})^{-1}$ , and  $(T)^{-1}$  denotes a numeric approximation of the inverse of  $T$ . The multiplication by  $J^{-1}$  is necessary to account for non-translational changes between the coordinate systems (Rao



et al., 2002).  $T_i^{p \rightarrow q}$  can be obtained from the transported motion fields using Eq. (7). This process can be repeated going in reverse from slice position  $q$  to  $p$  to arrive at the transformation  $T_j^{q \rightarrow p}$  at respiratory position  $j$ .

With approximate knowledge of the relations between the slice positions at arbitrary respiratory states an improved similarity measure can be defined as

$$\epsilon(X_i^{(p)}, X_j^{(q)}) = \frac{1}{2} \tilde{\mathcal{L}}_2(X_j^{(q)}, X_i^{(p)} \circ T_i^{p \rightarrow q}) + \frac{1}{2} \tilde{\mathcal{L}}_2(X_i^{(p)}, X_j^{(q)} \circ T_j^{q \rightarrow p}), \quad (9)$$

where, again  $\tilde{\mathcal{L}}_2$  denotes the normalised  $\mathcal{L}_2$ -distance, and  $I \circ T$  is the transformation of image  $I$  with the transformation  $T$ . This means the similarities were evaluated after transforming the slice at  $p$  into the coordinate system of  $q$  at the corresponding respiratory position, and vice versa. Using both the registrations  $T_i^{p \rightarrow q}$ , and  $T_j^{q \rightarrow p}$  improved the robustness of the similarity measure.

Note that the transformations  $T_i^{p \rightarrow q}$  and  $T_j^{q \rightarrow p}$  are only approximations, and might lose validity at deep inhale states where the slice relations change significantly with respect to an exhale state. However, even at those states the similarity measure was significantly improved over using the  $\mathcal{L}_2$ -distance directly.

The novel registration-based similarity kernel relating the slices at slice position  $p$  to the slices at slice position  $q$  can then be defined as

$$U_{ij}^{(p,q)} = K(\vec{X}_i^{(p)}, \vec{X}_j^{(q)}) = \exp \left( -\frac{\epsilon(X_i^{(p)}, X_j^{(q)})^2}{2\sigma^2} \right), \quad (10)$$

where again  $\sigma$  governs the kernel shape.

For the remainder of this paper SGA.REG will denote the simultaneous groupwise manifold alignment with the improved registration-based kernel.

### 3.2.3. Kernel Sparsification

To make the respiratory position matching more robust we used only a subset of the found similarities for the kernels in Eqs. (6) and (10), that is, the respective similarity kernels  $\mathbf{U}^{(p,q)}$  were sparsified. Note that in a simple k-nearest neighbour sparsification neighbourhood relations are directed and thus not symmetric, i.e. if ‘ $A$  is a neighbour of  $B$ ’ it does not necessarily follow that ‘ $B$  is a neighbour of  $A$ ’. This means that the sparsification will depend on which slice position the k-nearest neighbour operation is based on,

for example if we keep the  $k$ -nearest connections of slice position  $p$  to  $p+1$  we will get a different result than when looking for the  $k$ -nearest connections of slice position  $p+1$  to  $p$ . This in turn will lead to biased connectivities and will adversely affect the embedding. Furthermore, note that making the sparsified kernel symmetric by mirroring the connections is not permitted because the inserted connections do not correspond to any physical similarity, e.g. if slice  $i$  on slice position  $p$  is close to slice  $j$  on slice position  $q$  (i.e.  $\vec{X}_i^{(p)}$  is close to  $\vec{X}_j^{(q)}$ ) it does not follow that also  $\vec{X}_j^{(p)}$  is close to  $\vec{X}_i^{(q)}$ .

To overcome these limitations we used a kernel sparsification technique based on a global bipartite maximum edge similarity matching. That is, we calculated the matching in which every data point in  $\mathbf{X}_p$  was connected to exactly one data point in  $\mathbf{X}_q$ , and the sum of similarities over the corresponding edges  $U_{ij}$  was maximised. Figure 5 illustrates this process. The bipartite matching that maximises the similarity is highlighted in red in Figure 5b. This is equivalent to a combinatorial optimisation problem and can be solved using the Hungarian method (Kuhn, 1955). Note that here, if  $A$  is a neighbour of  $B$ , by definition  $B$  has to be a neighbour of  $A$ .

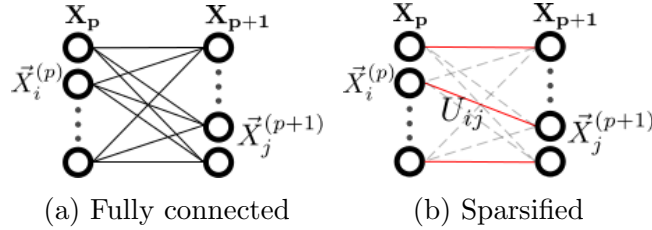


Figure 5: Graph sparsification of similarity kernel. (a) shows the fully connected graph, (b) shows the optimal one-to-one mapping.

The resulting graph can be written as a sparse matrix  $\mathbf{U}$ , which in every row and every column has exactly one non-zero entry,  $0 < U_{ij} \leq 1$ . This is similar to the case of labelled connections as in Ham et al. (2005), with the difference that here the labels are not known *a priori*, but instead have a certainty measure (i.e. the kernel value) attached to them.

### 3.3. Dynamic 3D Volume Reconstruction

After computing a low-dimensional embedding using either SGA.SIM or SGA.REG and thus establishing respiratory correspondences in the low-dimensional space as explained in Section 3.2, the only remaining step is to reconstruct the slices into a sequence of 3D volumes.

To this end we considered the first acquired slice, and looked up its closest neighbours in the low-dimensional space for all other slice positions in the aligned manifolds. The slices corresponding to those points were then stacked into a volume. This process was repeated for each slice  $X_i^{(p)}$  in the original acquisition order to arrive at a 3D+t reconstruction over the length of the acquisition. Note that because each slice position has been acquired 50 times (see Section 3.1) it can be assumed that most respiratory positions have been covered and there will be enough data to reconstruct a volume from each slice.

#### 4. Experiments

We evaluated the proposed simultaneous groupwise alignment methods with the similarity-based inter-data kernel (SGA.SIM) and the registration-based kernel (SGA.REG) against two other state-of-the-art slice stacking techniques, both of which have very similar acquisition schemes and reconstruction steps to our proposed method. They can be implemented directly in our framework by replacing the method used for establishing respiratory correspondences by alternative mechanisms (compare to Figure 2).

- Würslin et al. (2013) used a simultaneously acquired MR pencil beam navigator to establish respiratory correspondences between 2D MR slice-by-slice data. Thus each acquired slice has an associated navigator value. To reconstruct volumes, in the original work, the navigator values were used to bin the MR data from different slice positions into four gates. For the evaluation on our data we adapted this approach to allow for more accurate reconstructions. In the reconstruction step, for each slice a volume was reconstructed by selecting the data from all other slice positions with navigator values closest to the input slice. We will refer to this technique as PBNV.
- Dikaio et al. (2012) proposed establishing respiratory correspondences between 2D slice-by-slice MR data based solely on image similarities of neighbouring slices. For each slice the neighbouring slice that maximised the normalised mutual information was chosen. This process was repeated for the newly found slice until the whole volume was reconstructed. We will refer to this image-based method as IMBASED.

We performed three experiments. First, we evaluated all the techniques described above on realistic synthetic data generated with a known ground

truth. Secondly, we evaluated the methods on real data acquired using the acquisition protocol described in Section 3.1. Lastly, we performed a PET/MR simulation, in which we demonstrated the ability of the proposed method to generate robust motion fields in the lungs, which can be used to correct PET data for respiration induced motion.

#### 4.1. Experiment 1 - Validation on Synthetic Data

We performed experiments on synthetic data from 10 healthy volunteers. In order to generate the synthetic data we acquired a slice-by-slice breath hold volume at end-exhale covering all slice positions using the same acquisition sequence that we used for the real data, which was described in Section 3.1. In addition, for each volunteer, we acquired 50 low-resolution dynamic volumes on the same Phillips Achieva 3T MR system using a cardiac-triggered T1-weighted gradient echo sequence with an acquired image resolution of  $1.5 \times 4.1 \times 5 \text{ mm}^3$ , and an acquisition time of approximately 600 ms. An example of a coronal slice from such a volume was shown in Figure 1a. A 1D pencil beam navigator was recorded immediately before and after acquiring each dynamic image. To account for the length of the 3D acquisition the leading and trailing pencil beam navigators were averaged to arrive at a better estimate of respiratory position. Note that the pencil beam navigator was used only to implement the technique proposed by Würslin et al. (2013).

In the next step an exhale image was chosen from the 50 low-resolution dynamic images. This was used as a reference volume. We then obtained motion fields for each of the dynamic images by registering them to the reference volume using B-spline registration. Finally, we transformed the breath hold slices using the motion fields to arrive at synthetic slices at different respiratory positions. Note, however, that while the deformations of the thorax overall are realistic, the motion estimates inside the lungs are not reliable as they were derived from low-resolution volumes with little contrast in this area.

In order to evaluate the accuracy of the four reconstruction techniques we performed a leave-one-out (LOO) cross validation for each subject separately. For one of the synthetic slices the whole volume it belonged to was left out, apart from the slice itself. Next, a new volume was reconstructed around the slice from the remaining synthetic data, which resulted in an approximation  $\hat{\mathbf{V}}_{\text{out}}$  of the left-out ground-truth volume  $\mathbf{V}_{\text{out}}$ . The reconstruction error was estimated by calculating the  $\mathcal{L}_2$ -distance,  $\mathcal{L}_2(\hat{\mathbf{V}}_{\text{out}}, \mathbf{V}_{\text{out}})$ , between the two volumes. The process was repeated for each of the synthetic slices to get a

good estimate of the mean and the standard deviation of the reconstruction error.

In preliminary evaluations we found SGA.SIM and SGA.REG to be relatively insensitive to the choice of the parameters  $k$  (number of neighbours in the LLE cost function) and  $\sigma$  (the kernel shape parameter) as long as they were chosen from a reasonable range. We set the number of neighbours to half of the time points acquired at each slice position, i.e.  $k = 25$ , and  $\sigma = 0.5$  for the evaluation of both methods. We more closely investigated the significance of the weighting parameter  $\mu$  which governs the importance of the similarity kernel in Eq. (1) and the embedded dimensionality  $d$  of the manifold representation.

In order to estimate the inherent dimensionality of our slice-by-slice data we employed a technique developed for this purpose which employs fractal dimensions to estimate the underlying number of dimensions of high-dimensional data (Camastra and Vinciarelli, 2002). We applied the algorithm to our data from each slice position and each subject and formed an average. We found that the average inherent dimensionality of the data was close to 3 and we set the embedding dimensionality to  $d = 3$  accordingly.

We determined the optimal parameter value of  $\mu$  and the results of the LOO cross validation simultaneously in a 2-fold cross validation over all subjects. That is, we divided the 10 subjects into a tuning set of 5 and a test set of 5. We performed the LOO experiment on the tuning set for a range of different parameter values of  $\mu$ , combining the errors of the 5 subjects into a single error figure for each parameter. We chose a parameter range from 0.01 to 5 with a logarithmic spacing for this evaluation. Next, we selected the  $\mu$  with the minimum error and used it for the error evaluation for each of the subjects in the test set. Lastly, we swapped the positions of the test set and tuning set to evaluate the second fold. The optimum parameters for the two respective folds were  $\mu = \{0.144, 0.144\}$  for SGA.SIM and  $\mu = \{0.144, 0.224\}$  for SGA.REG.

Note that PBNV and IMBASED do not have any parameters that need to be tuned.

#### *4.2. Experiment 2 - Validation on Real Data*

In order to validate the proposed method on real data we acquired slice-by-slice data from 10 volunteers over the course of 24 to 28 minutes as described in Section 3.1. Note that all volunteers except volunteer H are the same as the volunteers from experiment 1. Volunteer H was not available for

the long scanning session and was replaced by volunteer K. Similar to the synthetic experiments, additionally a 1D pencil beam navigator was recorded for each slice. However, only a leading pencil beam navigator was acquired since acquiring both leading and trailing navigators was not possible on the scanner we used for an acquisition of this length.

For each of the examined slice stacking methods we reconstructed a volume around each of the acquired slices resulting in a time sequence of 3D volumes.

Since for the real data the ground truth is unknown we evaluated the consistency of reconstruction for each of the techniques. That is, for each acquired slice we reconstructed a volume  $\mathbf{V}$ , and for each slice  $s$  in this volume we reconstructed a new volume  $\hat{\mathbf{V}}_s$ . Ideally,  $\hat{\mathbf{V}}_s$  should be equal to  $\mathbf{V}$ . In practice, however, a different input slice position will give a different reconstruction. To estimate this reconstruction consistency we calculated the  $\mathcal{L}_2$ -distance,  $\mathcal{L}_2(\hat{\mathbf{V}}_s, \mathbf{V})$ , for each slice of each reconstructed volume. Consistency alone naturally does not give an indicator of the correctness of a method, as reconstructions can be consistently wrong. However, together with visual inspection of the results, and the results from the experiment on synthetic data from the previous section, the consistency error gives an indication as to how reproducible the shown results are.

Because for the real data we lacked a ground truth, there was no good measure to use for tuning the values of the parameters. Applying the dimensionality estimation technique to the real data in the same manner as before yielded an approximate average underlying dimensionality of 5 for each slice position. However, we found that neither SGA.SIM nor SGA.REG were robust to matching in such high dimensions. We suspect that this has to do with the fact that higher modes of variation might be slice position specific and not common to all slice positions as with the first three modes. Therefore, we used the same dimensionality as for the synthetic experiment, i.e.  $d = 3$  for both SGA.SIM and SGA.REG. Furthermore, we again used  $\sigma = 0.5$ ,  $k = 25$  and we fine tuned the value for  $\mu$  using visual inspection. We found that a value of  $\mu = 0.10$  worked well for both SGA.SIM and SGA.REG.

#### 4.3. Experiment 3 - PET/MR simulation

In the final experiment we applied IMBASED, SGA.SIM and SGA.REG for the retrospective respiratory motion correction of a realistic synthetic PET dataset using the reconstruct-transform-average (RTA) approach.

#### 4.3.1. Generation of Synthetic PET Data

Based on a breath hold slice-by-slice scan of volunteer A (shown in Figure 6a) we manually created a 3D fluorodeoxyglucose (FDG) uptake map with realistic standardised uptake values (SUVs) (Figure 6c). For the purpose of evaluating the accuracy of motion correction inside the lungs we additionally added an artificial lung tumour with 1.3 cm diameter and an SUV of 7 to the FDG maps. To simulate the deformations due to respiratory motion we transformed these maps using a set of *generating* motion fields. To obtain realistic motion estimates in the whole thorax, including the lung, the generating motion fields were derived from patient 2 of the publicly available POPI 4DCT dataset (Vandemeulebroucke et al., 2011) using B-spline registration and were transported to the coordinate system of our volunteer using the method described in Rao et al. (2002). A coronal slice of the CT volume transformed to the MR coordinate frame is shown in Figure 6b. Because the 10 available motion states in the POPI dataset were not enough for our simulation we linearly interpolated the motion fields to arrive at 30 (15 exhaling, 15 inhaling) motion states. The CT volume corresponding to end-exhale was then transformed by the 30 generating motion fields. These CT volumes served as attenuation maps in the following. Based on the 30 motion fields we simulated 30 gates of synthetic PET data ( $2 \times 2 \times 2$  mm<sup>3</sup> voxels, and 1.67 million coincidence events per gate) using the STIR package (Thielemans et al., 2006) with the ordered subsets expectation maximisation (OSEM) reconstruction algorithm (23 subsets, 11 iterations). The transformed attenuation maps were used to introduce attenuation effects into the PET simulations and also to correct for such effects in the reconstructions. A single gate of synthetic PET data at end-exhale is shown in Figure 6d. The tumour is indicated with an arrow. We also generated 30 PET gates without motion corruption which served as the ground truth for our evaluation. Lastly, we obtained 30 synthetic slice-by-slice MR volumes by transforming the breath hold slice-by-slice volume using the generating motion fields. This is the equivalent of simultaneously acquired slice-by-slice data in a real PET/MR scenario. Note that while the synthetic PET and MR data contains intra-cycle variabilities because inhaling and exhaling states were separated in the generating 4DCT data, it does not contain inter-cycle variabilities because in the 4DCT data each time point is an average of multiple breathing cycles.

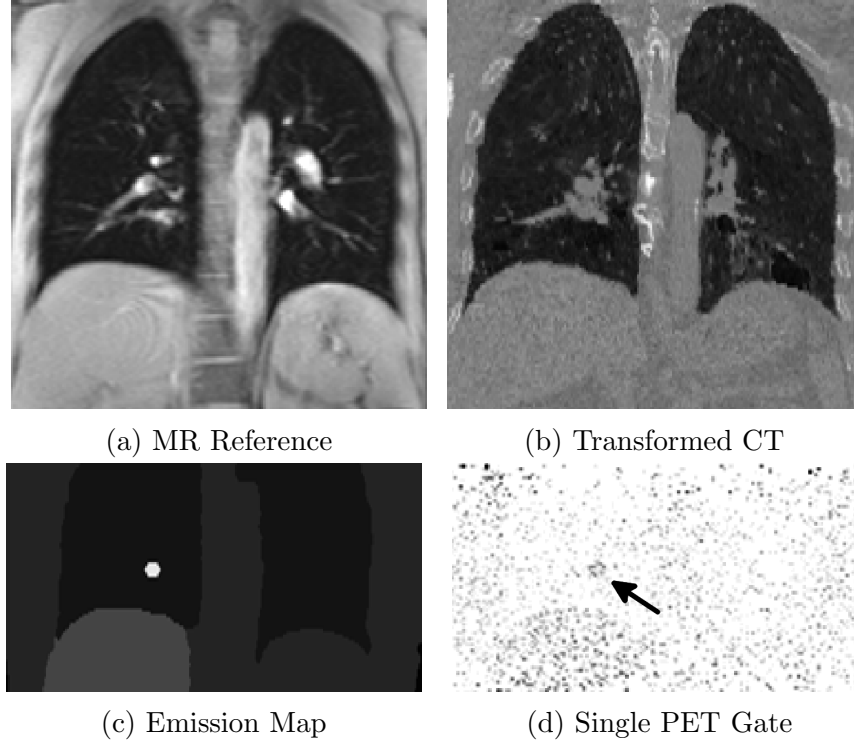


Figure 6: Coronal slices at the same A-P position through components of the synthetic PET simulation: (a) Reference MR scan for our volunteer, (b) exhale gate from POPI dataset warped to the coordinate system of our volunteer, (c) manually drawn emission map with an artificial tumour inserted in the lower lung (shown in white), and (d) an example of a single synthetic PET gate. The tumour is highlighted with an arrow.

#### 4.3.2. Retrospective Motion Correction

In the next step we reconstructed a volume from a synthetic slice for each time point using the IMBASED method, the SGA.SIM method and our proposed SGA.REG method. Again, the parameters for the latter two were chosen based on our findings on the synthetic data. The dimensionality was set to  $d = 3$  and the weight parameter was set to  $\mu = 0.144$ . Furthermore, we chose  $\sigma = 0.5$  and  $k = 15$ , which is again half of the time points acquired at each slice position. Note that performing this experiment using PBNV, by extracting an artificial pencil beam navigator signal from the images, would not be meaningful. This is because all of the slices originating from



one volume would have identical navigator values and the matching would always be perfect.

Similar to the acquisition of the real slice-by-slice data (see Section 3.1) we varied the input slice position for each time point with the square root of the total number of slice positions. For each of these reconstructed volumes we derived *backward* motion fields by registering the volumes derived using the three examined techniques to the MR reference volume using B-spline registration.

Finally, we transformed each of the synthetic PET gates back to the reference exhale state using the estimated backwards motion fields derived using each of the examined techniques, and averaged them to produce the final PET image. Additionally, we also averaged the PET gates without motion corruption to obtain a ground truth, and we reconstructed a PET volume without any motion correction. In order to compare the quality of motion correction we evaluated intensity profiles through the tumour for each of the reconstructions above.

## 5. Results

### 5.1. Experiment 1

The combined error scores on the synthetic tuning sets for each parameter value of  $\mu$  are shown in Figure 7. For comparison the errors obtained with the navigator based PBNV method on the tuning set are also shown. The errors obtained with IMBASED were much higher than the errors obtained using the other methods and were thus omitted on the plot. It can be seen that the optimal parameter  $\mu$  for both SGA.REG and SGA.SIM can be chosen from a range of values between approximately 0.05 and 0.5.

The results of the leave-one-out cross-validation on the respective synthetic data test sets are shown in Table 1.

Since the error distribution of all volunteers was symmetric but not normal, we evaluated the significance levels using a 1-tailed Wilcoxon signed rank test. We found that overall SGA.REG performed significantly better ( $p < 0.001$ ) than all other evaluated techniques, and SGA.SIM performed significantly better than PBNV and IMBASED. When examining results volunteer-by-volunteer we found that for 7 out of our 10 subjects the improvements of SGA.REG over the respective next best technique (highlighted in light grey in Table 1) were significant ( $p < 0.001$ ). For volunteers C, G and I,

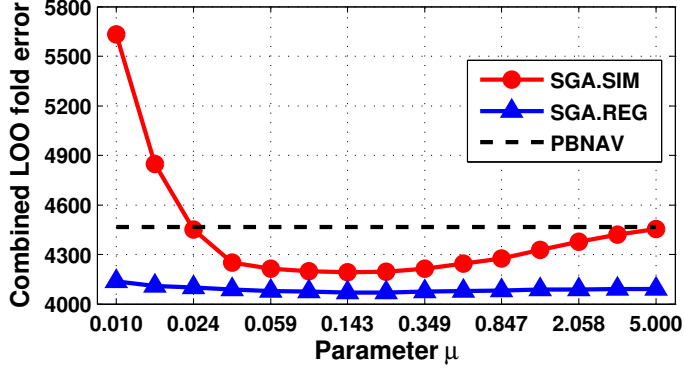


Figure 7: Combined fold error on the tuning set for SGA.SIM (red) and SGA.REG (blue). For comparison the average errors obtained with PBNV (black dotted line) on the tuning sets are also shown.

SGA.REG was not significantly better than SGA.SIM, but was significantly better than PBNV and IMBASED.

|             | PBNV<br>$\cdot 10^3$ | IMBASED<br>$\cdot 10^3$ | SGA.SIM<br>$\cdot 10^3$ | SGA.REG<br>$\cdot 10^3$ |
|-------------|----------------------|-------------------------|-------------------------|-------------------------|
| Volunteer A | $2.911 \pm 1.215$    | $4.604 \pm 2.150$       | $2.627 \pm 0.926$       | $2.508 \pm 0.902$       |
| Volunteer B | $3.940 \pm 1.113$    | $7.802 \pm 2.993$       | $3.818 \pm 0.955$       | $3.774 \pm 0.970$       |
| Volunteer C | $3.907 \pm 1.472$    | $4.822 \pm 1.543$       | $3.685 \pm 1.180$       | $3.649 \pm 1.161$       |
| Volunteer D | $3.183 \pm 1.031$    | $6.146 \pm 2.966$       | $3.478 \pm 1.909$       | $2.988 \pm 0.959$       |
| Volunteer E | $5.441 \pm 2.129$    | $7.995 \pm 2.773$       | $4.830 \pm 1.142$       | $4.673 \pm 1.200$       |
| Volunteer F | $7.076 \pm 2.650$    | $9.043 \pm 3.002$       | $6.192 \pm 1.717$       | $6.094 \pm 1.642$       |
| Volunteer G | $5.861 \pm 2.705$    | $5.941 \pm 2.041$       | $5.189 \pm 2.308$       | $5.105 \pm 2.247$       |
| Volunteer H | $4.184 \pm 1.698$    | $7.653 \pm 3.173$       | $3.557 \pm 1.042$       | $3.465 \pm 1.009$       |
| Volunteer I | $4.044 \pm 1.357$    | $6.723 \pm 2.282$       | $3.859 \pm 1.269$       | $3.806 \pm 1.271$       |
| Volunteer J | $3.967 \pm 1.294$    | $5.820 \pm 2.232$       | $3.460 \pm 0.851$       | $3.372 \pm 0.816$       |

Table 1: Means and standard deviations of the leave-one-out cross-validation experiment on synthetic slice-by-slice data for each volunteer. The method giving the lowest reconstruction error is highlighted in dark grey for each volunteer; the next best method is highlighted in light grey.

Examples of sagittal maximum intensity projections (MIPs) over volumes reconstructed for volunteer E using the four examined techniques, and their disparity with respect to the ground truth, are shown in Figure 8. The

magenta represents the ground truth, and the green shows structures of the reconstructed volume which do not appear in the ground truth. The top row (EXH) contains a typical end-exhale reconstruction and the bottom row (INH) a typical inhale reconstruction.

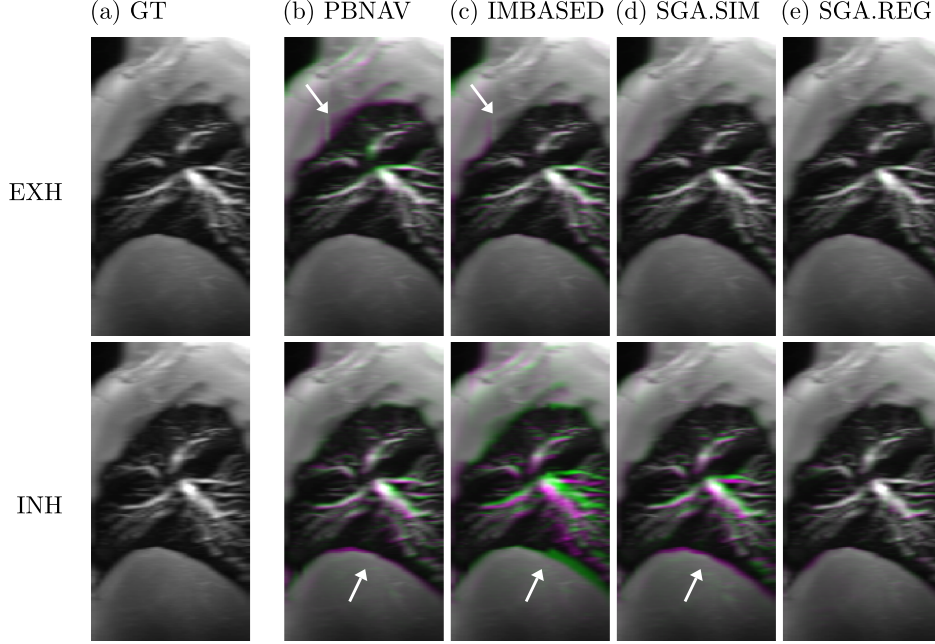


Figure 8: Examples of sagittal MIPs over the left lung of volunteer E obtained from synthetic slice-by-slice data reconstructed using the four examined techniques. The left most column (a), contains the ground truth (GT) MIP. The remaining columns show the disparity between the ground truth (shown in magenta) and the volumes reconstructed using the examined techniques: (b) PBNV, (c) IMBASED, (d) SGA.SIM, (e) SGA.REG (shown in green). The projections are shown at two respiratory positions: (EXH) end-exhale, and (INH) deep inhale. The white arrows indicate some of the disparities.

### 5.2. Experiment 2

In Figure 9 sagittal MIPs of volumes reconstructed from real slice-by-slice data using the different reconstruction techniques are presented. We show examples of one representative volunteer (volunteer B) at three distinct respiratory states (end-exhale, mid-inhale, and end-inhale). For the mid-inhale

and end-inhale reconstructions a small area inside the lungs was magnified to highlight the differences in vessel continuity.

For illustration the first two dimensions of the aligned embedding using SGA.REG of  $\mathcal{G}_2^{(18)}$  (group 18), which contains the manifolds of slice position 18 (blue), and 19 (red), is shown in Figure 10b. For comparison we also show the embeddings of the same two slice positions computed directly from LLE without any alignment in Figure 10a. The embedded low-dimensional coordinates which correspond to the slices used for slice positions 18, and 19 of the inhale (INH) reconstruction using SGA.REG in Figure 9 are highlighted with black squares in Figures 10a and b. In Figure 9 the two slice positions are indicated by small red triangles in the lower right.

The results of the consistency experiment are shown in Table 2. For all volunteers PBNV gave significantly ( $p < 0.001$ ) more consistent results than all other evaluated techniques, and SGA.REG gave significantly more consistent results than IMBASED and SGA.SIM. Again, significance was assessed using a 1-tailed Wilcoxon signed rank test as the error distributions were symmetric but not normal.

|             | PBNV<br>·10 <sup>3</sup> | IMBASED<br>·10 <sup>3</sup> | SGA.SIM<br>·10 <sup>3</sup> | SGA.REG<br>·10 <sup>3</sup> |
|-------------|--------------------------|-----------------------------|-----------------------------|-----------------------------|
| Volunteer A | 1.777 ± 1.632            | 5.730 ± 3.671               | 3.240 ± 2.203               | 2.738 ± 2.072               |
| Volunteer B | 2.181 ± 1.842            | 6.960 ± 4.278               | 4.117 ± 2.873               | 3.554 ± 2.309               |
| Volunteer C | 2.850 ± 2.160            | 7.807 ± 4.327               | 4.860 ± 3.382               | 4.347 ± 2.887               |
| Volunteer D | 2.566 ± 2.513            | 8.717 ± 5.700               | 4.654 ± 3.452               | 4.301 ± 3.289               |
| Volunteer E | 2.585 ± 2.321            | 8.123 ± 5.137               | 5.034 ± 3.264               | 4.282 ± 2.808               |
| Volunteer F | 4.026 ± 2.959            | 8.560 ± 5.140               | 6.268 ± 3.980               | 6.079 ± 3.860               |
| Volunteer G | 4.166 ± 2.942            | 8.598 ± 4.447               | 6.417 ± 4.150               | 5.550 ± 3.681               |
| Volunteer I | 2.807 ± 1.995            | 8.309 ± 5.256               | 5.416 ± 3.571               | 4.082 ± 2.573               |
| Volunteer J | 2.766 ± 2.069            | 6.712 ± 3.933               | 4.658 ± 3.234               | 3.889 ± 2.459               |
| Volunteer K | 3.841 ± 2.600            | 8.636 ± 5.528               | 6.083 ± 3.795               | 4.693 ± 3.035               |

Table 2: Means and standard deviations of the consistency experiment on real slice-by-slice data for each volunteer. The method giving the lowest consistency error for each volunteer is highlighted in dark grey, and the method with the second to lowest consistency error is highlighted in light grey.

### 5.3. Experiment 3

The results of the PET motion correction experiment are shown in Figures 11 and 12. A coronal slice including the tumour through the ground truth reconstruction without motion corruption is shown in Figure 11a and the reconstruction without motion correction is shown in Figure 11b. The motion corrected reconstructions using IMBASED, SGA.SIM, and SGA.REG are shown in Figures 11c, 11d, and 11e. Close-up views of the tumour for (a)-(e) are shown in Figure 11f. Finally, the line profiles through the tumour along the line indicated in Figure 11a, are shown in Figure 12.

## 6. Discussion

We have presented a novel method for the navigator-less, accurate reconstruction of high-resolution, high-contrast dynamic 3D volumes from long slice-by-slice MR acquisitions based on the simultaneous groupwise manifold alignment of neighbouring slice positions (SGA.REG). Such volumes have many potential applications. Here, we applied them for the motion correction of simultaneously acquired PET data. The present work is an extension of the SGA.SIM technique which we previously published (Baumgartner et al., 2013) by a registration-based similarity kernel to account for anatomical differences between adjacent slice positions.

The proposed technique captures the breathing motion of each slice position in a low-dimensional manifold embedding. For a sufficiently high embedded dimensionality, this representation of motion has the potential to model all respiratory inter-cycle as well as intra-cycle variabilities such as amplitude variations, baseline shifts and hysteresis in the manifold representation. Such variabilities are then taken into account during the reconstruction resulting in more accurate volumes than current methods. In future work we plan to investigate in more detail to which degree the different types of variations contribute to this improvement.

We compared our technique against two recently proposed PET motion correction schemes based on MR slice-by-slice data: the purely image-based technique (IMBASED) proposed by Dikaïos et al. (2012), and the pencil beam navigator-based technique (PBNV) proposed by Würslin et al. (2013).

On synthetic data (experiment 1) the proposed SGA.REG method significantly outperformed all other methods in terms of reconstruction accuracy (see Table 1). The results were confirmed by looking at maximum intensity

projections in the sagittal plane of the synthetic data (Figure 8). At end-exhale all methods tended to perform very well, most likely because exhale states are generally more reproducible over the course of a long acquisition than inhale states. However, especially the reconstructions using PBNV suffered from artefacts in the anterior chest-lung interface (indicated by arrows in the top row of Figure 8), which can be explained by intra-cycle or inter-cycle breathing variabilities that cannot be captured by a technique based on a 1D navigator signal. The biggest improvements of SGA.REG over the other techniques were achieved for inhale states, such as the one shown in the bottom row of Figure 8, where breathing variabilities are largest. We observed that IMBASED did not perform well in our experiments and often gave implausible reconstructions such as the one in the bottom row of Figure 8.

We made similar observations in our experiments on real data (experiment 2). By visual inspection of the results we found that exhale states can be reconstructed accurately with most techniques, as can be seen in the through-plane maximum intensity projections of reconstructions in Figure 9. Only IMBASED sometimes suffered from implausible reconstructions such as the kink in the diaphragm indicated by an arrow in the EXH row. For other respiratory positions the effects of misalignment became more evident. In mid-inhale frames such as the one shown in the MID row of Figure 9 we occasionally observed inaccuracies in the anterior chest-lung interface for PBNV (highlighted by an arrow) similar to the ones observed on synthetic data. These kind of artefacts are not surprising as PBNV derives all its respiratory information from the diaphragm and doesn't have information about other areas. The vessel close-up views show that SGA.REG gives the smoothest estimate in this area of all examined techniques. Artefacts like the ones mentioned above became much worse at inhale states, such as the one shown in the bottom row of Figure 9 where respiratory variabilities are the largest. In this frame PBNV suffers from large artefacts on the diaphragm and the anterior chest-lung interface as highlighted by arrows, which are likely due to respiratory intra- and inter-cycle variations. Also the vessels are distorted and intermittent for PBNV while SGA.REG retains smooth continuous vessels. As can be seen in the same figure, IMBASED and SGA.SIM give anatomically implausible reconstructions at inhale. SGA.REG is the only technique which incorporates knowledge of the anatomical ground truth in the form of our novel registration-based inter-dataset kernel for matching the slices. We attribute the improvements of SGA.REG over SGA.SIM and

IMBASED to this fact.

The consistency results of experiment 2 shown in Table 2 allow to measure how consistently the visual results shown in Figure 9 can be reproduced over all time frames. The consistency results should however not be mistaken as a measure for reconstruction accuracy and should not be interpreted on their own. PBNV gave the most consistent results for all volunteers. This can be understood when considering that a 1D signal is easier to match consistently to 1D signals from other slice positions. This, however, comes at the cost of not being able to model inter- and intra-cycle breathing variabilities, which cannot be captured by a 1D signal only. The other techniques require matchings in higher dimensions which have more ambiguity, but allow for such variabilities to be captured. The high consistency of PBNV also means that errors due mismatches caused by breathing variabilities were also consistently reproduced. SGA.REG exhibited a high consistency for all volunteers when compared to SGA.SIM and IMBASED and offered a good compromise between reconstruction reproducibility and ability to model complex breathing patterns.

Acquiring data using a slice-by-slice acquisition protocol as described in Section 3.1 allowed excellent image contrast and made it possible to image vessel structures inside the lungs which cannot be visualised using a dynamic 3D MR acquisition protocol. In the synthetic PET/MR experiment (experiment 3) these additional structures allowed us to extract very detailed and reliable motion fields of the whole region of interest in the thorax. By incorporating these into an RTA motion correction scheme we could successfully correct realistic, synthetic PET data for respiratory motion. Note that because of the way the data was generated in this experiment, only intra-cycle variabilities could be modelled. Also, in this experiment, we used the generating motion fields to transform the attenuation maps used in the reconstructions, which would obviously not be possible when using real data. In practice, these maps could be transformed using motion fields estimated from the images produced using our technique. As is shown in Figure 11, the reconstructions using SGA.SIM (Figure 11d) and SGA.REG (Figure 11e) both very closely match the ground truth with no motion corruption (Figure 11a). The reconstruction using IMBASED (Figure 11c) is markedly less defined. This is confirmed by looking at the line profiles through the tumour shown in Figure 12. The profile through the SGA.REG reconstruction is very similar to the motionless ground truth. The profile from the SGA.SIM reconstruction is slightly less defined.

It has been shown by Polycarpou et al. (2011) that under some circumstances MCIR may provide superior PET motion correction than RTA, but comes at greater computational cost. Our technique can also be used to provide motion estimates for MCIR-based reconstructions, and in future work we plan to investigate this.

During a PET/MR session typically many different MR scans need to be acquired for, amongst other things, attenuation correction and visualisation of different aspects of the anatomy. Note that the MR-based PET motion correction scheme proposed here could be applied with minimal scanning overheads, since acquiring one 2D slice for 160 *ms* per cardiac cycle is sufficient to retrospectively obtain high-contrast 3D volumes for the entire duration of a PET imaging session. In the remainder of the time the scanner could be used for other MR imaging relevant to the PET session. In the future it would also be conceivable to integrate part of those additional scanning requirements into our method, for example, by temporally interleaving partial MR data acquisitions using different protocols and reconstructing them retrospectively using SGA.REG.

For this work we acquired 50 dynamic slices of each slice position which resulted in acquisition times of 24-28 minutes. This number was chosen arbitrarily and the minimum number of dynamics required to build reliable reconstructions will be investigated in future work. For smaller datasets interpolating amongst the *k*-nearest neighbours on the manifold as was briefly discussed in Section 2.3 may be beneficial.

The largest deformations due to breathing motion are in the superior-inferior (S-I) and A-P directions. Therefore, the choice of coronal slices may be suboptimal and it may be preferable to use sagittal slices since these two deformations are within the sagittal plane. The proposed SGA.REG method is currently not robust when applied to sagittal slices, because the anatomy undergoes much larger changes from slice position to slice position in this plane. Extending the method to arbitrary slice orientations will be the subject of future investigations. Nevertheless, through-plane motion is only slightly hampering the performance of our method because the changes in slice appearance due to A-P motion are also captured in the manifold.

The proposed technique also has potential application in MR-guided treatments such as MR-guided high-intensity focused ultrasound (HIFU) (Köhler et al., 2011). In a calibration phase 2D slices could be acquired and embedded in an aligned low-dimensional representation as described in this work. The aligned embeddings contain information about all respiratory



states that are likely to occur and can thus be viewed as a motion model (McClelland et al., 2013). In a subsequent application phase a volume could be generated from previously unseen data by acquiring a new slice at a convenient slice position, embedding it into the aligned manifold embedding and looking up the closest slices at all other slice positions. In MR-guided HIFU this could be applied for fast online updating of guidance information. In a similar fashion the proposed method could also have potential applications in the guidance of radiotherapy treatments in a hybrid MR linear accelerator setup such as recently proposed by Stam et al. (2012).

## 7. Conclusion

The recent emergence of hybrid PET/MR systems is a promising technology for many applications which currently use PET/CT systems. Acquiring MR images during a PET procedure comes at no additional radiation exposure to the patient and thus opens up possibilities for non-ionising and very accurate respiratory motion correction. Using 1D navigators for gating or reconstructing MR volumes suffers from inaccuracies due to breathing variabilities and does not use the full potential of MR imaging. Here we have proposed an effective and accurate motion correction technique, which does not rely on navigators, has the potential to image breathing variabilities, and may be combined easily with other MR scanning requirements that may arise during a PET/MR imaging session.

## Acknowledgements

This work was funded by EPSRC programme grant EP/H046410/1. This research was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy’s and St Thomas’ NHS Foundation Trust and King’s College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. This work was partly funded by the EU FP7 SUBLIMA project. We like to thank Irene Polycarpou for various helpful discussions.

## References

Bai, W., Brady, M., 2009. Regularized B-spline deformable registration for respiratory motion correction in PET images. *Phys Med Biol.* 54, 2719.

- Baumgartner, C., Kolbitsch, C., McClelland, J., Rueckert, D., King, A., 2013. Groupwise simultaneous manifold alignment for high-resolution dynamic MR imaging of respiratory motion, in: Proc. IPMI, Springer. pp. 232–243.
- Belkin, M., Niyogi, P., 2003. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* 15, 1373–1396.
- Beyer, T., Townsend, D., Brun, T., Kinahan, P., Charron, M., Roddy, R., Jerin, J., Young, J., Byars, L., Nutt, R., et al., 2000. A combined PET/CT scanner for clinical oncology. *J Nucl Med.* 41, 1369–1379.
- Bhatia, K., Rao, A., Price, A., Wolz, R., Hajnal, J., Rueckert, D., 2012. Hierarchical manifold learning, in: Proc. MICCAI.. Springer, pp. 512–519.
- Blackall, J., Ahmad, S., Miquel, M., McClelland, J., Landau, D., Hawkes, D., 2006. MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. *Phys Med Biol.* 51, 4147–4169.
- Camasta, F., Vinciarelli, A., 2002. Estimating the intrinsic dimension of data with a fractal-based method. *IEEE T Pattern Anal.* 24, 1404–1407.
- Dawood, M., Lang, N., Jiang, X., Schafers, K., 2006. Lung motion correction on respiratory gated 3-D PET/CT images. *IEEE T Med Imaging.* 25, 476–485.
- Dikaio, N., Izquierdo-Garcia, D., Graves, M., Mani, V., Fayad, Z., Fryer, T., 2012. MRI-based motion correction of thoracic PET: initial comparison of acquisition protocols and correction strategies suitable for simultaneous PET/MRI systems. *Eur radiol.* 22, 439–446.
- Ding, K., Yin, Y., Cao, K., Christensen, G., Lin, C.L., Hoffman, E., Reinhardt, J., 2009. Evaluation of lobar biomechanics during respiration using image registration, in: Proc. MICCAI.. Springer, pp. 739–746.
- Ehman, R., Felmlee, J., 1989. Adaptive technique for high-definition MR imaging of moving structures. *Radiology* 173, 255–263.
- Georg, M., Souvenir, R., Hope, A., Pless, R., 2008. Manifold learning for 4D CT reconstruction of the lung, in: Proc. IEEE CVPRW., pp. 1–8.

- Ham, J., Lee, D., Saul, L., 2005. Semisupervised alignment of manifolds, in: AI and Statistics, pp. 120–127.
- King, A., Boubertakh, R., Rhode, K., Ma, Y., Chinchapatnam, P., Gao, G., Tangcharoen, T., Ginks, M., Cooklin, M., Gill, J., et al., 2009. A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures. *Med Image Anal.* 13, 419–431.
- King, A., Buerger, C., Tsoumpas, C., Marsden, P., Schaeffter, T., 2012. Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator. *Med Image Anal.* 16, 252–264.
- Klein, G., Reutter, B., Ho, M., Reed, J., Huesman, R., 1998. Real-time system for respiratory-cardiac gating in positron tomography. *IEEE T Nucl Sci.* 45, 2139–2143.
- Köhler, M., Denis de Senneville, B., Quesson, B., Moonen, C., Ries, M., 2011. Spectrally selective pencil-beam navigator for motion compensation of MR-guided high-intensity focused ultrasound therapy of abdominal organs. *Magn Reson Med.* 66, 102–111.
- Kuhn, H., 1955. The Hungarian method for the assignment problem. *Nav Res Logist Q.* 2, 83–97.
- McClelland, J., Hawkes, D., Schaeffter, T., King, A., 2013. Respiratory motion models: A review. *Med Image Anal.* 17, 19 – 42.
- McClelland, J., Hughes, S., Modat, M., Qureshi, A., Ahmad, S., Landau, D., Ourselin, S., Hawkes, D., 2011. Inter-fraction variations in respiratory motion models. *Phys Med Biol.* 56, 251–272.
- Nehmeh, S., Erdi, Y., Ling, C., Rosenzweig, K., Squire, O., Braban, L., Ford, E., Sidhu, K., Mageras, G., Larson, S., et al., 2002. Effect of respiratory gating on reducing lung motion artifacts in PET imaging of lung cancer. *Med Phys.* 43, 876–881.
- Polycarpou, I., Tsoumpas, C., Marsden, P., 2011. Statistical evaluation of PET motion correction methods using MR derived motion fields, in: IEEE Nucl Sci Conf R., IEEE. pp. 3579–3585.

- Qiao, F., Pan, T., Clark Jr., J.W., Mawlawi, O., 2006. A motion-incorporated reconstruction method for gated PET studies. *Phys Med Biol.* 51, 3769.
- Rao, A., Sanchez-Ortiz, G., Chandrashekar, R., Lorenzo-Valdés, M., Mohiaddin, R., Rueckert, D., 2002. Comparison of cardiac motion across subjects using non-rigid registration, in: *Proc. MICCAI.* Springer, pp. 722–729.
- Robson, M., Gatehouse, P., Bydder, M., Bydder, G., 2003. Magnetic resonance: an introduction to ultrashort TE (UTE) imaging. *J Comput Assist Tomo.* 27, 825–846.
- Roweis, S., Saul, L., 2000. Nonlinear dimensionality reduction by locally linear embedding. *Science* 290, 2323–2326.
- Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D., 1999. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE T Med Imaging.* 18, 712–721.
- Seppenwoolde, Y., Shirato, H., Kitamura, K., Shimizu, S., van Herk, M., Lebesque, J., Miyasaka, K., 2002. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol.* 53, 822–834.
- Stam, M., Crijns, S., Zonnenberg, B., Barendrecht, M., van Vulpen, M., Lagendijk, J., Raaymakers, B., 2012. Navigators for motion detection during real-time MRI-guided radiotherapy. *Phys Med Biol.* 57, 6797–6805.
- Tenenbaum, J., De Silva, V., Langford, J., 2000. A global geometric framework for nonlinear dimensionality reduction. *Science* 290, 2319–2323.
- Thielemans, K., Mustafovic, S., Tsoumpas, C., 2006. STIR: software for tomographic image reconstruction release 2, in: *IEEE Nucl Sci Conf R., IEEE.* pp. 2174–2176.
- Thielemans, K., Rathore, S., Engbrant, F., Razifar, P., 2011. Device-less gating for PET/CT using PCA, in: *Proc. IEEE NSS/MIC., IEEE.* pp. 3904–3910.
- Torki, M., Elgammal, A., Lee, C., 2010. Learning a joint manifold representation from multiple data sets, in: *Proc. IEEE ICPR.,* pp. 1068–1071.

- Tsoumpas, C., Mackewn, J., Halsted, P., King, A., Buerger, C., Totman, J., Schaeffter, T., Marsden, P., 2010. Simultaneous PET–MR acquisition and MR-derived motion fields for correction of non-rigid motion in PET. *Ann Nucl Med.* 24, 745–750.
- Vandemeulebroucke, J., Rit, S., Kybic, J., Clarysse, P., Sarrut, D., 2011. Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. *Med Phys.* 38, 166–181.
- von Siebenthal, M., Székely, G., Gamper, U., Boesiger, P., Lomax, A., Cattin, P., 2007a. 4D MR imaging of respiratory organ motion and its variability. *Phys Med Biol.* 52, 1547–1564.
- von Siebenthal, M., Székely, G., Lomax, A., Cattin, P., 2007b. Systematic errors in respiratory gating due to intrafraction deformations of the liver. *Med Phys.* 34, 3620–3629.
- Wachinger, C., Yigitsoy, M., Rijkhorst, E., Navab, N., 2011. Manifold learning for image-based breathing gating in ultrasound and MRI. *Med Image Anal.* 16, 806–818.
- Wang, C., Mahadevan, S., 2008. Manifold alignment using Procrustes analysis, in: *Proc. ICML*.
- Würslin, C., Schmidt, H., Martirosian, P., Brendle, C., Boss, A., Schwenzer, N., Stegger, L., 2013. Respiratory motion correction in oncologic PET using T1-weighted MR imaging on a simultaneous whole-body PET/MR system. *J Nucl Med.* 54, 464–471.
- Zaidi, H., Del Guerra, A., 2011. An outlook on future design of hybrid PET/MRI systems. *Med Phys.* 38, 5667–5689.
- Zhai, D., Li, B., Chang, H., Shan, S., Chen, X., Gao, W., 2010. Manifold alignment via corresponding projections, in: *Proc. BMVC.*, pp. 3–11.

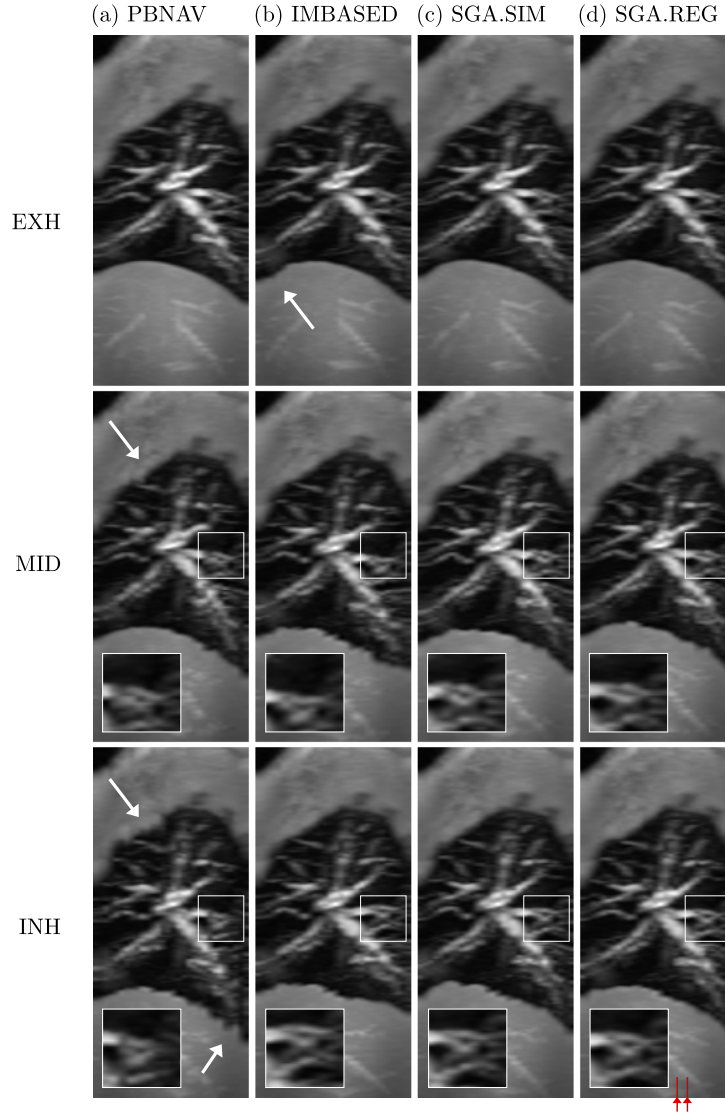


Figure 9: Examples of sagittal MIPs over the left lung of volunteer B obtained from real slice-by-slice data reconstructed using the four examined methods: (a) PBNV, (b) IMBASED, (c) SGA.SIM, and (d) SGA.REG. The projections are shown at three respiratory positions: (EXH) end-exhale, (MID) mid-inhale, (INH) at end-inhale. Magnifications of the square area indicated with white boxes in the middle and bottom rows are shown in the lower left corner of the respective images. The red markers in the lower right mark the slice positions shown in the sample manifold embedding in Figure 10.

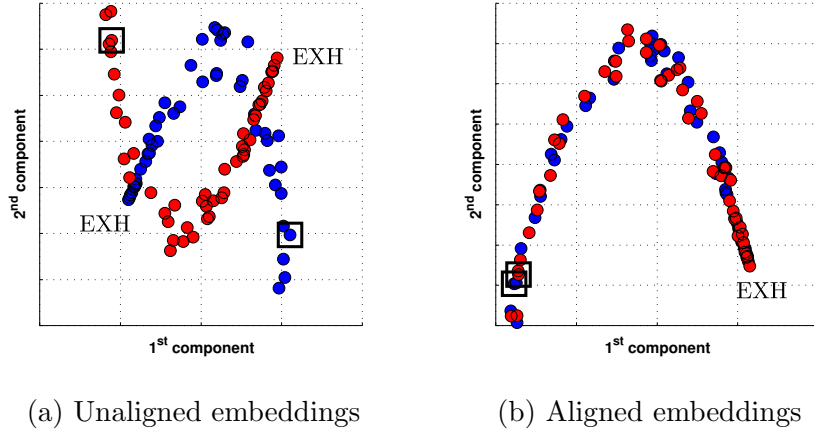


Figure 10: Unaligned and aligned embeddings for slice positions 18 (blue) and 19 (red) of volunteer B. (a) shows the first two dimensions of the 3-dimensional embeddings of the two slice positions as obtained directly from LLE without alignment. (b) shows the first two dimensions of group 18 of the aligned SGA.REG embedding. The manifold coordinates of the slices that were chosen for the inhale reconstruction of SGA.REG in Figure 9 are highlighted with black rectangles.

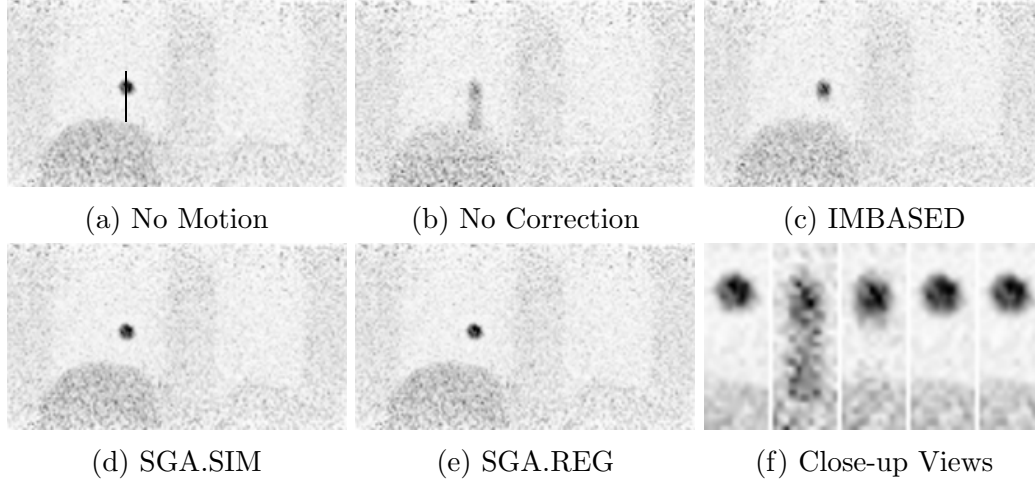


Figure 11: Result of motion correction applied to the synthetic PET data: (a) The sum of the PET gates with no motion (ground truth), (b) PET reconstruction without motion correction, and motion corrected PET reconstructions using (c) IMBASED, (d) SGA.SIM and SGA.REG. Finally close-up views of the tumour for (a)-(e) are shown in (f). The line profiles in Figure 12 were calculated along the dotted line in (a).

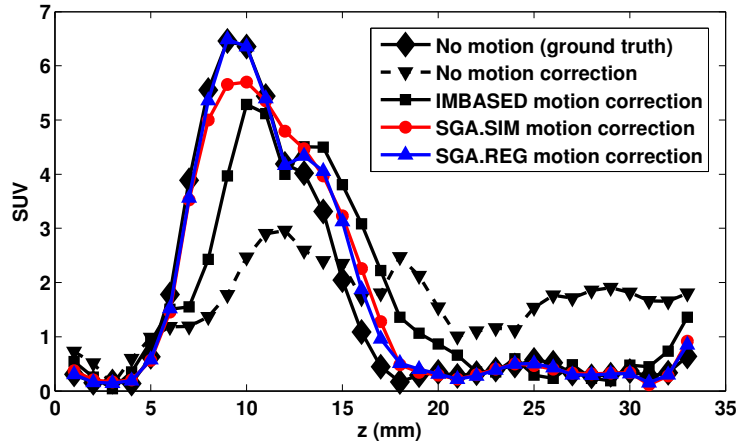


Figure 12: SUV line profiles through the tumour along the line indicated in Figure 11a for the motionless ground truth, no motion correction, and motion correction using IMBASED, SGA.SIM and SGA.REG. The motionless ground truth and the SGA.REG motion correction have almost identical profiles between  $z = 5 \text{ mm}$  and  $z = 12 \text{ mm}$ .